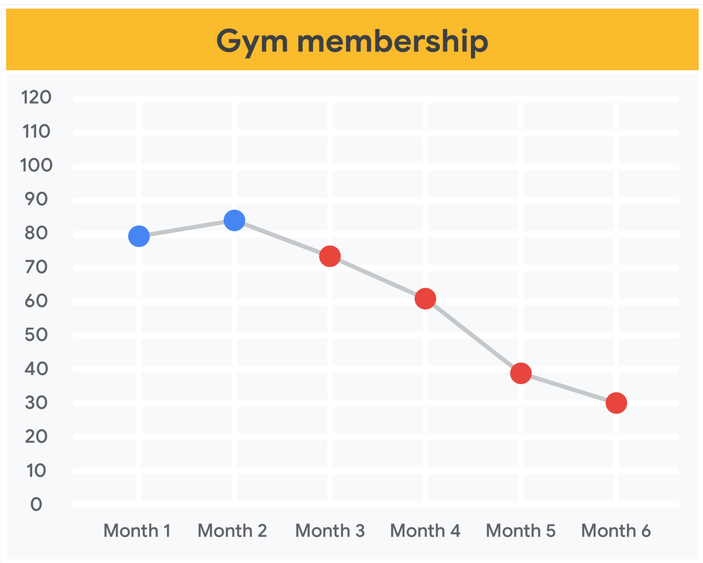
The six steps of data analysis

In an earlier video, you learned that **data analysis** is the process of collecting and organizing information to help draw conclusions, solve problems, make informed decisions, and support your goals. In this reading, we will go over the key parts of the data analysis process.

There are six main steps involved in data analysis: **Ask, prepare, process, analyze, share** and **act**. Let's break these down one by one.

During the **Ask** phase, ask key questions to help frame your analysis, starting with: What is the problem? When defining the problem, look at the current state of the business and identify how it is different from the ideal state. Usually, there is an obstacle in the way or something wrong that needs to be fixed. At this stage, you want to be as specific as possible. You also want to stay focused on the problem itself, not just the symptoms. For example, imagine you are doing data analysis for a gym that is losing memberships. You could ask: Why do we keep losing members? But a better and more specific question would be: What factors are negatively impacting the member experience? That way, when you set off to do your research, you know exactly what to look for.



Another part of the **Ask** stage is identifying your stakeholders and understanding their expectations. There can be lots of stakeholders on a project, and each of them can make decisions, influence actions, and weigh in on strategies. Each stakeholder will also have specific goals they want to meet. It is pretty common for a stakeholder to come to you with a problem that needs solving. But before you begin your analysis, you need to be clear about what they are asking of you. For example, if your manager assigns you a project related to analyzing the gym's business risk, it would be a good idea to confirm whether they want you to analyze all types of risks that could affect the gym or just risks related to weather or seasonal trends.

After you have a clear direction, it is time to move to the **Prepare** stage. This is where you collect and store the data you will use for the upcoming analysis process.

Let's turn back to our gym membership example. To collect data on the member experience, you decide to send surveys to the gym's members asking for feedback about their experience. To make sure you get specific answers, you ask them to offer feedback in three distinct categories: upkeep of the facility, customer service, and membership cost**.** You also leave room for them to write in a response. When you get the member surveys back, it is important that you have an organized system for tracking and filing them.

This stage is when it is time to **Process** your data. In this step, you will "clean" your data, which means you will enter your data into a spreadsheet, or another tool of your choice, and eliminate any inconsistencies and inaccuracies that can get in the way of results. While collecting data, be sure to get rid of any duplicate responses or biased data. This helps you know that any decisions made from the analysis are based on facts and that they are fair and unbiased. For example, if you noticed duplicate responses from a single gym member when sorting through the surveys, you would need to get rid of the copies to be sure your data set is accurate.

During this stage, it is also important to check the data you prepared to make sure it is complete and correct and that there are no typos or other errors.

Now it is time to **Analyze**. In this stage, you take a close look at your data to draw conclusions, make predictions, and decide on next steps. Here, you will transform and organize the data in a way that highlights the full scope of the results so you can figure out what it all means. You can create visualizations using charts and graphs to determine if there are any trends or patterns within the data or any need for additional research.

In our gym membership example, let's say you notice 50% of the members wrote in an additional response on the survey citing that the equipment is outdated. The survey also showed that 75% of the responses cited "expensive membership fees." When looking at the 50% of responses citing "outdated equipment" and 75% of responses citing "expensive membership fees" side by side on a graph, you may be able to deduce that these responses inform one another. Members feel like the experience just isn't worth the price. You might conclude that the gym should invest in new equipment if they want to keep members and add value to the membership fee.

Once you have asked questions to figure out the problem—then prepared, processed, and analyzed the data—it is time to **Share** your findings. In this stage, you use **data visualization** to organize your data in a format that is clear and digestible for your audience. When sharing, you can offer the insights you gained during your analysis to help stakeholders make effective, data-driven decisions for solving the problem.

And finally, you are ready to **Act**! In the final stage of your data analysis, the business takes all of the insights you have provided and puts them into action to solve the original business problem.

Conducting a data analysis is an essential process for understanding a business' needs and challenges and determining effective solutions. These six foundational steps—**ask, prepare, process, analyze, share,** and **act**—will help set you up for success!